

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-012092

(43)Date of publication of application : 21.01.1994

(51)Int.Cl. G10L 3/00
G10L 3/00

(21)Application number : 05-098442

(71)Applicant : AMERICAN TELEPH & TELEGR CO <ATT>

(22)Date of filing : 02.04.1993

(72)Inventor : BOCCHIERI ENRICO L
GOKGEN SEDAT I
MIKKILINENI RAJENDRA P
ROE DAVID B
WILPON JAY GORDON

(30)Priority

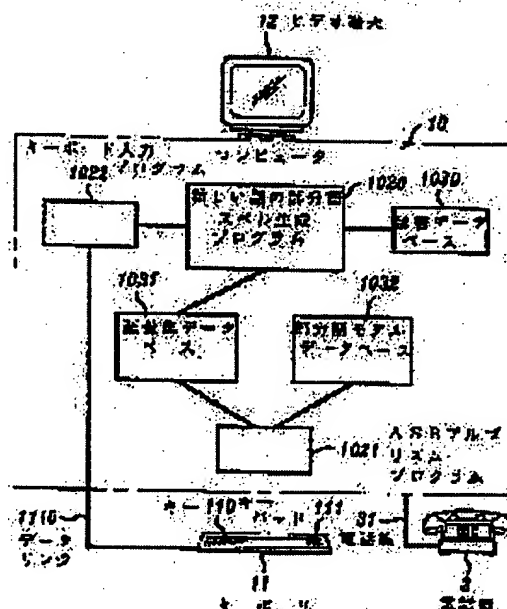
Priority number : 92 862394 ...Priority date : 02.04.1992 ...Priority country : US

(54) SPEECH RECOGNITION DEVICE AND ITS OPERATING METHOD

(57)Abstract:

PURPOSE: To make it possible to input information used by a user's business application by comparing a constituted partial speech model with speech notations recorded in a vocabulary data base in response to the reception of speaking data.

CONSTITUTION: After spoken alphanumeric data is received through a telephone line 31, an automatic speech recognition (ASR) algorithm program 1021 accesses a partial word model data base 1032 and constitutes a model which consists of a series of phonemes and represents the sound of the received spoken words. Then the ASR algorithm program 1021 accesses the vocabulary data base 1031 and compares the constituted series phoneme model with voice representation series recorded in the vocabulary data base 1031. When the constituted series phoneme model matches a voice representation series recorded in the vocabulary data base 1031, it is recognized as a recorded word corresponding to the matching voice representation series after the vocalization.



LEGAL STATUS

[Date of request for examination] 16.01.1996

[Date of sending the examiner's decision of rejection] 15.02.2000

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平6-12092

(43)公開日 平成6年(1994)1月21日

(51)Int.Cl.⁵

G10L 3/00

識別記号

531 D 7627-5H

561 G 7627-5H

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数20(全 12 頁)

(21)出願番号 特願平5-98442

(22)出願日 平成5年(1993)4月2日

(31)優先権主張番号 862394

(32)優先日 1992年4月2日

(33)優先権主張国 米国(US)

(71)出願人 390035493

アメリカン テレフォン アンド テレグ
ラフ カムパニー

AMERICAN TELEPHONE
AND TELEGRAPH COMPA
NY

アメリカ合衆国 10013-2412 ニューヨ
ーク ニューヨーク アヴェニュー オブ
ジ アメリカズ 32

(74)代理人 弁理士 三俣 弘文

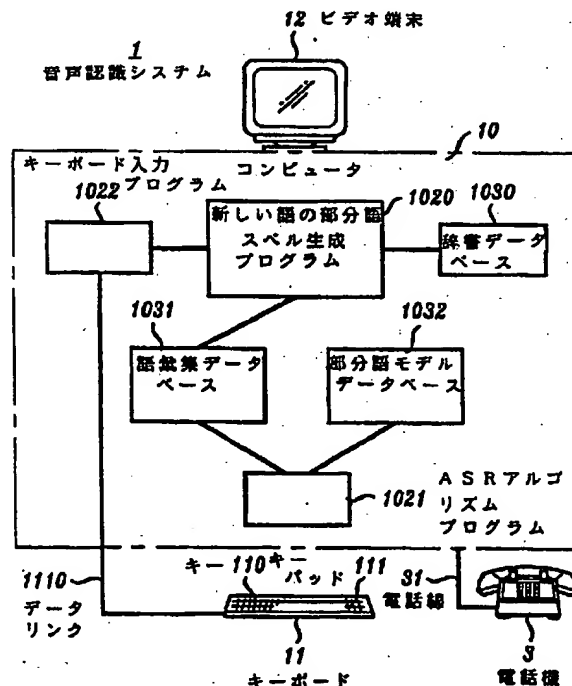
最終頁に続く

(54)【発明の名称】 音声認識装置およびその動作方法

(57)【要約】

【目的】 音声認識システムにデータを記録し、記録データに対応する発話データを認識する装置および方法を実現する。

【構成】 入力されたデータに回答して、音声表示の列を生成する。入力データおよび生成した音声表示列が、音声認識システムの語彙集に記録される。発話データの受信に回答して、その発話データを特徴づける部分語のモデルを構成し、構成した部分語列に一致する音声表示列に対応するデータとして発話データを認識するために、構成した部分語モデルを語彙集に記録された音声表示列と比較する。



【特許請求の範囲】

【請求項1】 語から音声表示の列を生成し、その語および生成した音声表示列を語彙集に記録するプロセッサ手段からなり、

このプロセッサ手段が、

発話語の受信に応答して、その発話語を特徴づける部分語のモデルを構成し、構成した部分語モデルと生成した音声表示列が一致した場合に発話語を前記語として認識するために、構成した部分語モデルを語彙集に記録された音声表示列と比較することを特徴とする音声認識装置。

【請求項2】 プロセッサ手段が、語および各格納語の音声表示を格納する辞書データベースを有することを特徴とする請求項1の装置。

【請求項3】 プロセッサ手段が、前記語およびその変形を、前記語およびその語変形を表現する音声表示の列とともに記録する語彙集データベースを有することを特徴とする請求項2の装置。

【請求項4】 プロセッサ手段が、前記音声認識装置への語の入力に応答して、辞書データベースにアクセスし、入力された語およびその組合せによって識別される音声表示の列を辞書データベースの音声表示から生成し、入力された語および生成した音声表示列を語彙集データベースに記録する手段を有することを特徴とする請求項3の装置。

【請求項5】 プロセッサ手段が、発話を区別するために利用される音声の小単位をそれぞれ識別する部分語を格納する部分語モデルデータベースを有することを特徴とする請求項4の装置。

【請求項6】 プロセッサ手段が、前記音声認識装置による語の受信に応答して、部分語モデルデータベースにアクセスし、前記受信した発話語を表現する部分語の列からなるモデルを構成することを特徴とする請求項5の装置。

【請求項7】 アクセス・構成手段が、語彙集データベースにアクセスし、構成した部分語モデルを、語彙集データベースに記録された音声表示列と比較し、構成した列部分語モデルと一致する音声表示列とともに記録された語として前記発話語を認識する手段を有することを特徴とする請求項6の装置。

【請求項8】 語から音声表示の列を生成し、その語および生成した音声表示列を語彙集データベースに記録する手段と、

発話語の受信に応答して、その発話語の音の特徴づける音素の列のモデルを構成し、語彙集データベースにアクセスし、モデル音素列と語彙集データベースに記録された音声表示列のうちの1つが一致した場合に、一致した語彙集データベースに記録された音声表示列とともに記録された語として前記発話語を認識するために、構成した音素列モデルを語彙集データベースに記録された音声

表示列と比較する手段とからなることを特徴とする音声認識装置。

【請求項9】 語および各格納語の音声表示を格納する辞書データベースと、
前記語およびその変形を、前記語およびその語変形を表現する音声表示の列とともに記録する語彙集データベースと、
音声認識装置への語の入力に応答して、辞書データベースにアクセスし、入力された語およびその変形によって識別される音声表示の列を辞書データベースから生成し、入力された語および生成した音声表示列を語彙集データベースに記録する第1プロセッサ手段と、
発話を区別するために利用される音声の小単位をそれぞれ識別する音素を格納する部分語モデルデータベースと、

前記音声認識装置による発話語の受信に応答して、部分語モデルデータベースにアクセスし、前記受信した発話語を表現する音声の小単位を表す音素の列からなるモデルを構成し、構成した音素モデル列と一致する音声表示列とともに記録された語として前記発話語を認識するために、構成した音素モデル列を語彙集データベースに記録された音声表示列と比較する第2プロセッサ手段とからなることを特徴とする音声認識装置。

【請求項10】 語の変形をそれぞれ表現する音声表示を格納し、発話を区別するために利用される音声の小単位をそれぞれ識別する音素を格納する手段と、
語を識別する音声表示の列を生成し、その語および生成した音声表示列の両方を前記格納手段に記録するプロセッサ手段とからなり、このプロセッサ手段が、
発話語の受信に応答して、その発話語を特徴づける音素の列のモデルを構成し、モデル音素列が記録された音声表示列のうちの1つと一致した場合に、記録された音声表示列に対応しそれとともに記録された語として前記発話語を認識するために、構成した音素列モデルを記録された音声表示列と比較することを特徴とする音声認識装置。

【請求項11】 語から音声表示の列を生成し、その語および生成した音声表示列を語彙集に記録するステップと、

発話語の受信後、その発話語を特徴づける音素の列のモデルを構成し、発話語を語彙集に記録された語のうちの1つとして認識するために、構成した音素列モデルを語彙集に記録された音声表示列と比較するステップとからなることを特徴とする音声認識装置の動作方法。

【請求項12】 前記生成・記録ステップが、語および各格納語の音声表示を辞書データベースに格納するステップを有することを特徴とする請求項11の方法。

【請求項13】 前記生成・記録ステップが、前記語およびその変形を、前記語およびその語変形を表現する音声表示の列とともに語彙集データベースに記録するステ

3

ップをさらに有することを特徴とする請求項12の方法。

【請求項14】 前記生成・記録ステップが、前記音声認識装置への語の入力にตอบสนองして、辞書データベースにアクセスするステップと、入力された語およびその組合せによって識別される音声表示の列を辞書データベースに格納された音声表示から生成するステップと、入力された語および生成した音声表示列を語彙集データベースに記録するステップとをさらに有することを特徴とする請求項13の方法。

【請求項15】 発話を区別するために利用される音声の小単位をそれぞれ識別する音素を部分語モデルデータベースに格納するステップをさらに有することを特徴とする請求項14の方法。

【請求項16】 前記構成・比較ステップが、前記音声認識装置による語の受信にตอบสนองして、部分語モデルデータベースにアクセスするステップと、前記受信した発話を表現する音素の列からなるモデルを構成するステップとを有することを特徴とする請求項15の方法。

【請求項17】 前記構成・比較ステップが、語彙集データベースにアクセスするステップと、構成した列音素モデルを、語彙集データベースに記録された音声表示列と比較するステップと、構成した列音素モデルと一致する音声表示列とともに記録された語として前記発話を認識するステップとをさらに有することを特徴とする請求項16の方法。

【請求項18】 プロセッサを有する音声認識装置において、このプロセッサは、このプロセッサと結合されたキーボードで入力された語にตอบสนองして、キーボード入力語から音声表示の列を生成し、その語および生成した音声表示列を語彙集に記録し、発話語の受信にตอบสนองして、その発話語を特徴づける音素の列モデルを構成し、その発話語をキーボード入力語として認識するために、構成した音素モデル列を語彙集に記録された音声表示列と比較することを特徴とする音声認識装置。

【請求項19】 英数字データの変形をそれぞれ表現する音声表示を格納し、発話を区別するために利用される音声の小単位をそれぞれ識別する音素を格納する手段と、英数字データを識別する音声表示の列を生成し、その英数字データおよびその英数字データに対応して生成した音声表示列の両方を前記格納手段に記録する手段とからなり、この手段が、発話英数字データの受信にตอบสนองして、その発話英数字データの特徴づける音素の列のモデルを構成し、モデル音素列が記録された音声表示列のうちの1つと一致した場合に、記録された音声表示列に対応する英数字データと

4

して前記発話英数字データを認識するために、構成した音素列モデルを記録された音声表示列と比較することを特徴とする音声認識装置。

【請求項20】 英数字データから音声表示の列を計算し、その英数字データおよびそれに対応する生成した音声表示列の両方を語彙集に記録するステップと、入力された英数字データの受信後、入力された英数字データの特徴づける音素の列のモデルを構成し、入力された英数字データを語彙集に記録された英数字データのうちの1つとして認識するために、構成した音素列モデルを語彙集に記録された音声表示列と比較するステップとからなることを特徴とする音声認識方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声認識システムで利用される単語モデルを生成する装置および方法に関し、特に、生成された単語を認識するために使用される単語モデルに関する。

【0002】

【従来の技術】 音声認識システムは通常、さまざまなアプリケーションで使用される完全な数詞、語および句を表すデータベース辞書記憶パターン（例えば音声テンプレートまたはモデル）を有する。

【0003】

【発明が解決しようとする課題】 システム所有者およびその所有者の顧客のみによって頻繁に使用され、それらの者に固有の音声テンプレートのデータベース辞書を、個々の音声認識システムが必要とする際に問題が生じる。例えば、病院の音声認識システムは、青果卸売業者の辞書に記録された音声テンプレートとは異なる別個の音声テンプレートの辞書を必要とする。すなわち、各音声認識システムは、それぞれ特定の音声認識システムアプリケーションに必要なテンプレートを構成するために、多数の音声データサンプルの収集およびその音声データサンプルの手作業の確認を必要とする。従って、多数の音声データサンプルの収集および確認を必要とせず、システムの所有者が自己のビジネスアプリケーションで利用される情報を音声認識システムに入力することができるような、音声認識システムのための自動音声認識器、および、その自動音声認識器を動作させる方法が必要である。

【0004】

【課題を解決するための手段】 上記の問題は、音声認識システムの語彙集データベース内の音声認識システムにデータを入力し、それによってシステムが、発話データを格納されたデータとして認識することを可能にする装置および方法によって解決される。プロセッサが、入力されたデータから音声表記を生成し、そのデータ（例えば英数字）および生成された対応する音声表記の両方を語彙集データベースに記録・格納する。プロセッサは、

5

発話データの部分語特性の列からなるモデルを構成することによって、発話データの受信にตอบสนองし、構成した部分語モデルを、語彙集データベースに記録された音声表記と比較する。構成された部分語モデルが、記録されたデータに対応する記録された音声表記と一致すると、発話データは記録されたデータとして認識される。

【0005】

【実施例】本発明の実施例において、音声認識システム1(図1)は、発話データ(例えば英数字データ)を認識するように配置される。一般的に、このようなシステムはさまざまなアプリケーション(例えば、大規模店の受注部)において使用可能である。このようなシステムは、例えば、電話線(例えば電話線31)に接続され、電話機3を使用する発呼顧客に音声指示メッセージを送信するために使用される。

【0006】音声認識システム1は、店に注文する顧客によって発話された英数字データを認識することによって、発呼顧客にตอบสนองする。他のアプリケーションでは、音声認識システム1は、電話交換手アプリケーションにおいて、発呼者発話電話番号を認識するため、および、株式仲買業者によって、顧客の注文および情報要求を認識するために使用される。さらに他のアプリケーションでは、音声認識システム1の原理は、コンピュータが、コンピュータのユーザによって発話される英数字データを認識することを可能にするために、コンピュータとともに使用される。

【0007】音声認識システム1は、例えば、AT&T 386または486コンピュータのようなコンピュータ10である。このコンピュータ(図2)は、本発明の理解のためには詳細は不要であり、一般的に、中央処理装置101、メモリユニット102、およびインタフェースユニット100を有し、これらはそれぞれアドレス線、データ線および制御線によってデータバス104に接続される。

【0008】インタフェースユニット100は、データ入力装置(例えばキーボード11)と相互接続されたデータリンク1110をデータバス104と結合し、キーボード11に入力された英数字データ(例えば語、数および情報)が中央処理装置101およびメモリユニット102と交換されるようにする。データバス104はまた、アドレス線、データ線および制御線によってデータベース構造体(例えばデータベース構造体1030、1031および1032)に格納・記録するために使用されるデータ記憶装置103とも相互接続される。音声認識システム1として使用されるコンピュータ10(図1)は、データおよび情報を表示するために使用されるビデオ端末12を有することも可能である。

【0009】図2で、本発明の原理によれば、動作時には、プログラム(例えば、新しい語の部分語スペル生成プログラム1020、自動音声認識(ASR)アルゴリ

6

ズムプログラム1021およびキーボード入力プログラム1022)は、メモリユニット102に格納され、音声認識システム1の動作を制御する適当な時刻に中央処理装置101にロードされる。データベース構造体(辞書データベース1030、語彙集データベース1031、および部分語モデルデータベース1032)はデータ記憶装置103に格納され、情報の読み出しおよび格納のために中央処理装置103によってアクセスされる。

【0010】辞書データベース1030(図3)は、語および各格納語の音声表示を格納する。一般的に、語は標準的な辞書に存在する語であり、格納語の音声表示とともに辞書データベース1030に格納される。または、語は、新しい語の部分語スペル生成プログラム1020によって計算される。例えば、語「house」は、対応する音声表示「haus」によって表現される。同様に、格納語「car」は、対応する音声表示「kaar」によって表現され、語「computer」は、対応する音声表示「kempyuter」によって表現される。さらに、「d」および「w」のような文字は、表示「dly」および「dahbixlyw」によって表現される。

【0011】ASR語彙集データベース1031は、語および記録される語の変形を表現する音声表示の文字列とともに、語およびその変形を記録・格納する。典型的な例では、新しい語の部分語スペル生成プログラム1020は、語「car」および対応する音声表示文字列「kaar」を語彙集データベース1031に記録する。他の例では、語「rail」および「road」の変形が、「railroad」として、対応する音声表示文字列「reylrowd」とともに記録される。

【0012】部分語モデルデータベース1032は、発話を区別するために利用する音声の小単位を識別する各音素を格納する。例えば、「cot」の中にあるような音を表す音素「aa」が部分語モデルデータベース1032に格納される。同様に、データベース1032は、「buf」、「dad」および「bait」の中にあるような音を表す各音素「ah」、「d」、「ey」を格納する。

【0013】図1で、語もしくは数字またはその列は、キーボード11のキー110およびキーパッド111上の語または数字をタイプすることによって音声認識システム1に入力される。他のアプリケーションでは、新たな追加語および数字は、テープ、ディスクドライブ、外部メモリまたは、プログラムによってコンピュータ10にデータを入力するいくつかの異なる周知の手段(例えばキーボード入力プログラム1022)によって語および数字を入力することによって音声認識システム1に追加される。

【0014】新しい語の部分語スペル生成プログラム1

020(図2)は、本発明の実施例によれば、中央処理装置101の動作を制御し、辞書データベース1030にアクセスすることによって語および数字の入力に应答する。図3で、入力された語(例えば「house」)が発見された場合、新しい語の部分語スペル生成プログラム1020は、入力語によって識別される辞書データの辞書音声表示から音声表示列「haus」を生成する。

【0015】辞書データベース1030に格納された語の変形(例えば語「railroad」)が音声認識システム1に入力された場合、新しい語の部分語スペル生成プログラム1020は、辞書データベース1030にアクセスし、入力された「railroad」によって識別される語「rail」および「road」とともに格納された辞書データベース1030音声表示「rey l」および「row d」から音声表示列「reyl row d」を生成する。

【0016】新しい語の部分語スペル生成プログラム1020は、各入力語およびそれに対応して生成された音声表示列を語彙集データベース1031に記録する。こうして、入力語「car」および「railroad」ならびにこれらに対応する音声列「kaar」および「reyl row d」が、音声認識システム1によって受信される発話語を認識する際に使用するために、語彙集データベース1031に記録される。

【0017】図1で、電話線31を通じて発話英数字データを受信後、自動音声認識(ASR)アルゴリズムプログラム1021(図3)は、部分語モデルデータベース1032にアクセスし、音素(発話を区別するために利用される音声の小単位をそれぞれ識別する)の列からなり、受信した発話語の音を表現するモデルを構成する。こうして、発話語「car」が受信された場合、ASRアルゴリズムプログラム1021は、部分語モデルデータベース1032にアクセスし、受信した発話語「car」を識別する音素列「k aa r」として音素「k」、「aa」および「r」を構成する。同様に、ASRアルゴリズムプログラム1021は、発話語「railroad」の受信に应答して、部分語モデルデータベース1032にアクセスし、受信した発話語「railroad」を表現する音素の列「r ey l r ow d」からなるモデルを構成する。

【0018】受信した発話語を表現する列音素モデルを構成した後、ASRアルゴリズムプログラム1021は語彙集データベース1031にアクセスし、構成された列音素モデルを、語彙集データベース1031に記録された音声表示列と比較する。構成された列音素モデルが、語彙集データベース1031に記録された音声表示列と一致した場合、発話語は、一致した音声表示列に対応する記録された語として認識される。

【0019】例えば、音声認識システム1(図1)が

発話語「railroad」を受信した場合、ASRアルゴリズムプログラム1021(図3)は、部分語モデルデータベース1032にアクセスし、列音素モデル「r ey l r ow d」を構成する。次に、プログラム1021は語彙集データベース1031にアクセスし、構成された列音素モデル「r ey l r ow d」を、語彙集に記録された音声表示列と比較する。構成された列音素モデル「r ey l row d」が記録された音声表示列「reyl row d」と一致した場合、発話語「railroad」は音声表示列「reyl row d」とともに記録された語「railroad」として認識される。

【0020】音声認識システム1(図1)は、通常、辞書データベース1030および語彙集データベース1031にそれぞれ多数の語を格納し記録している。システムが追加語を認識することを可能にするために、追加語が、音声認識システム1に入力され、語彙集データベース1031に記録される。このような語は、音声認識システム1の特定アプリケーションで使用される固有語であることが可能である。本発明の実施例では、特定アプリケーションに固有の特殊な語「dwd」が音声認識システム1に入力されたと仮定する。

【0021】「dwd」のような発話語を認識するために音声認識システム1を動作させる方法は、語「dwd」から音声表示列「diydahbixlyuwdiy」を生成するステップと、語「dwd」および生成された音声表示列「diydahbixlyuwdiy」の両方を語彙集データベース1031(図3)に記録するステップからなる。発話語「dwd」を受信した後、本発明の音声認識システム1を動作させる方法は、部分語モデルデータベース1032にアクセスし、発話語「dwd」の音を特徴づける音素のモデル列「diydahbixlyuwdiy」を構成する。

【0022】構成された音素列モデル「diydahbixlyuwdiy」は、語彙集に記録された音声表示列と比較され、構成された音素列モデル「diydahbixlyuwdiy」と語彙集データベース1031に記録された音声表示列「diydahbixlyuwdiy」の一致があった場合、発話語は、一致した音声表示列「diydahbixlyuwdiy」とともに記録された語「dwd」として認識される。

【0023】図1で、本発明の動作方法において、標準語の辞書および各標準語の音声表示が最初に辞書データベース1030に格納される。同様に、部分語モデルデータベース1032は、最初に、「d」および「w」のような音素(発話を区別するために利用される音声の小単位をそれぞれ識別する)を格納する。新しい語(例えば語「dwd」)を音声認識システム1に入力するため

には、語「dwd」がキーボード11にタイプされ、キーボード入力プログラム1022に入力される(図4、ステップ102200、102201)。

【0024】新しい語の部分語スペル生成プログラム1020(図1)は、英数字データおよびその変形(例えば語「dwd」)の入力にตอบสนองして、辞書データベース1030にアクセスする(図4、ステップ102000、102001)。入力語が辞書データベース1030に発見された場合(ステップ102002)、新しい語の部分語スペル生成プログラム1020は、その後および対応する音声表示列を選択し(ステップ102003)、その後および対応する音声表示列の両方を語彙集データベース1031に記録する(ステップ102005、102006)。

【0025】入力語(例えば語変形「dwd」)が辞書データベース1030に発見されなかった場合(ステップ102002)、新しい語の部分語スペル生成プログラム1020は、図3で、成分語「d」および「w」を辞書データベース1030から選択し、音声表示列「d i y d a h b i x l y u w d i y」を生成する(ステップ102004)。語変形「dwd」および生成された音声表示列「d i y d a h b i x l y u w d i y」は、語彙集データベース1031に記録される(ステップ102005、102006)。

【0026】さらに多くの語が音声認識システム1に入力されている場合(ステップ102007)、ステップ102201およびステップ102000~102007が、すべての入力語が語彙集データベース1031に記録されるまで反復され、その後、中央処理装置101は、新しい語の部分語スペル生成プログラム1020を終了する(ステップ102008)。

【0027】ASRアルゴリズムプログラム1021に従って動作する音声認識システム1(図1)は、電話線31を通じて受信された発話語(例えば語「dwd」)にตอบสนองして、部分語モデルデータベース1032にアクセスする(図5、ステップ102100、102101、102102)。発話語「dwd」の受信にตอบสนองして、ASRアルゴリズムプログラム1021は、受信した発話語「dwd」を表現する音素の列からなるモデルの構成を開始する(ステップ102103)。

【0028】このモデルの構成は、受信した「dwd」の音に従い、音「d」の結果、列モデルの最初の部分で「d i y」が選択される(ステップ102104)。モデルは完成していない(ステップ102105)ため、ステップ102101~102105が反復され、「w」の音の結果、音素「d a h b i x l y u w」が部分語モデルデータベース1032から選択されモデルに付加される。最後の音「d」の結果として音素「d i y」が列に付加され、発話語「dwd」を表現する音素列モデル「d i y d a h b i x

l y u w d i y」が完成する(ステップ102105)。

【0029】発話語「dwd」を表現する構成された音素列モデル「d i y d a h b i x l y u w d i y」の完成後、ASRアルゴリズムプログラム1021は、語彙集データベース1031にアクセスし、構成された列音素モデル「d i y d a h b i x l y u w d i y」を、語彙集データベースに記録された音声表示列と比較する(ステップ102106、102107、102108)。

【0030】構成された音素列モデル「d i y d a h b i x l y u w d i y」と、記録された音声表示列「d i y d a h b i x l y u w d i y」の一致があった場合(ステップ102109)、発話語「dwd」が、語彙集データベース1031に記録された語「dwd」として認識される(ステップ102110)。さらに発話語がある場合(ステップ102111)、すべての語が認識されるまでステップ102101~102111が反復される。

【0031】一致が失敗した場合(ステップ102109)、かつ、一致試行が所定試行数以下である場合(ステップ102113)、ASRアルゴリズムプログラム1021は、反復試行数を増加させ、発呼者に、発話語を反復するよう促す(ステップ102114)。この場合、ステップ102101~102109が反復される。試行数が所定試行数を超過した場合、エラーが発生したと判断される(ステップ102115)。

【0032】一致プロセス(ステップ102109)中に、信頼性認識因子が、構成された音素列モデルと、語彙集データベース1031に記録された音声表示列の一致に対して割り当てられる。割り当てられた信頼性認識因子が所定しきい値を超過した場合、構成された音素列モデルに対応する発話語が、語彙集データベース1031に記録されている、一致した音声表示列に対応する語として認識されることが決定される。

【0033】

【発明の効果】以上述べたごとく、本発明によれば、多数の音声データサンプルの収集および確認を必要とせず、システムの所有者が自己のビジネスアプリケーションで使用する情報を音声認識システムに入力することができる、音声認識システムのための自動音声認識器、および、その自動音声認識器を動作させる方法が実現される。

【図面の簡単な説明】

【図1】本発明の原理を実現する音声認識システムのブロック図である。

【図2】本発明の原理による、図1に示された音声認識システムのプロセッサの詳細のブロック図である。

【図3】図1および2に示す音声認識システムとともに使用されるプロセッサプログラムおよびデータベースフ

11

12

アイルの例の図である。

【図4】本発明の原理による、図1に示された音声認識システムの動作の流れ図である。

【図5】本発明の原理による、図1に示された音声認識システムの動作の流れ図である。

【符号の説明】

1 音声認識システム

3 電話機

10 コンピュータ

11 キーボード

12 ビデオ端末

31 電話線

100 インタフェースユニット

101 中央処理装置

102 メモリユニット

103 データ記憶装置

104 データバス

110 キー

111 キーパッド

1020 新しい語の部分語スペル生成プログラム

1021 自動音声認識 (ASR) アルゴリズムプログラム

1022 キーボード入力プログラム

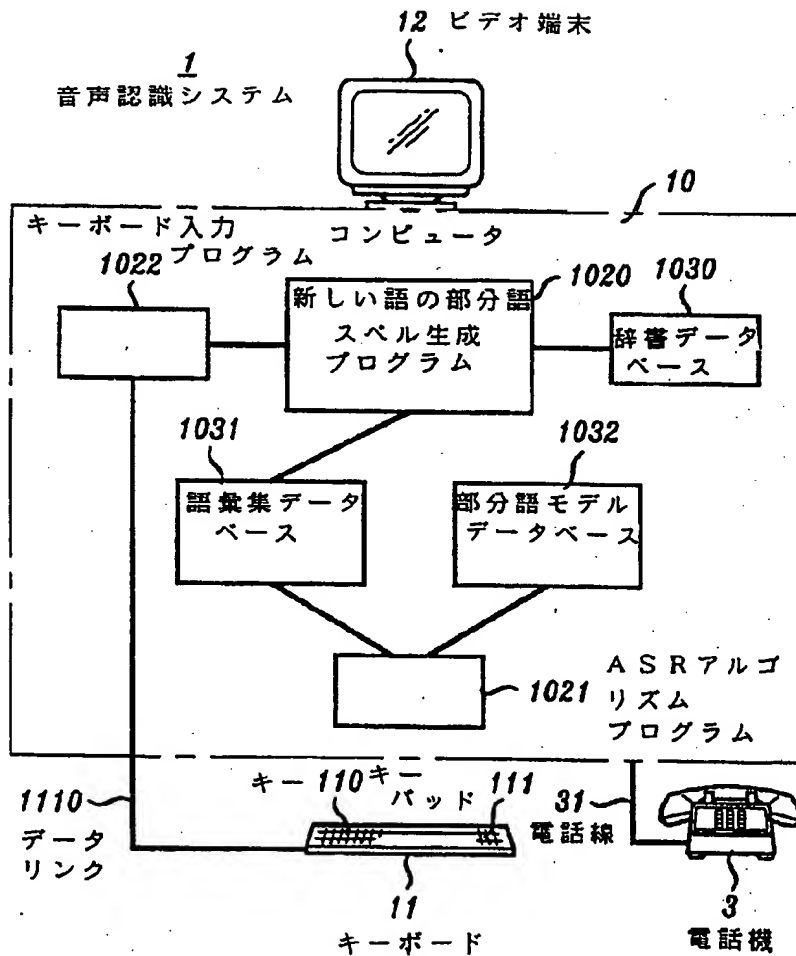
10 1030 辞書データベース

1031 語彙集データベース

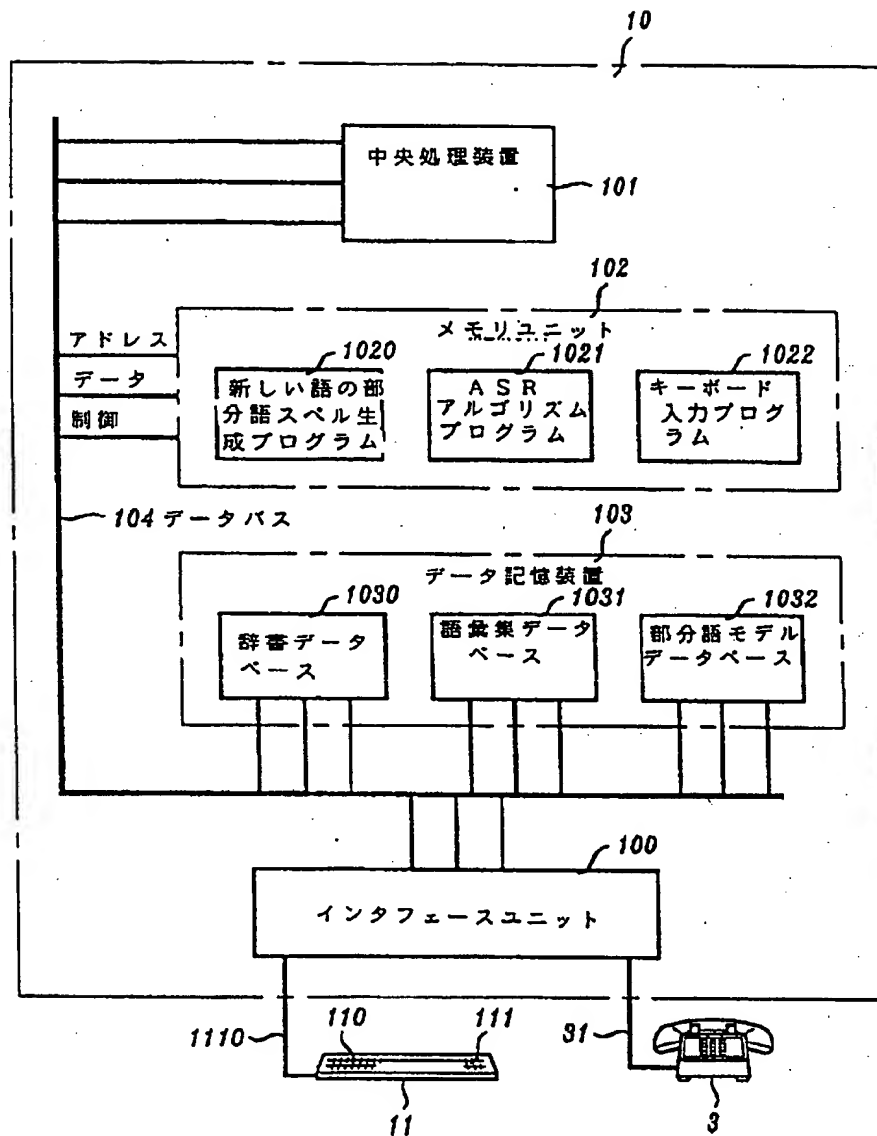
1032 部分語モデルデータベース

1110 データリンク

【図1】

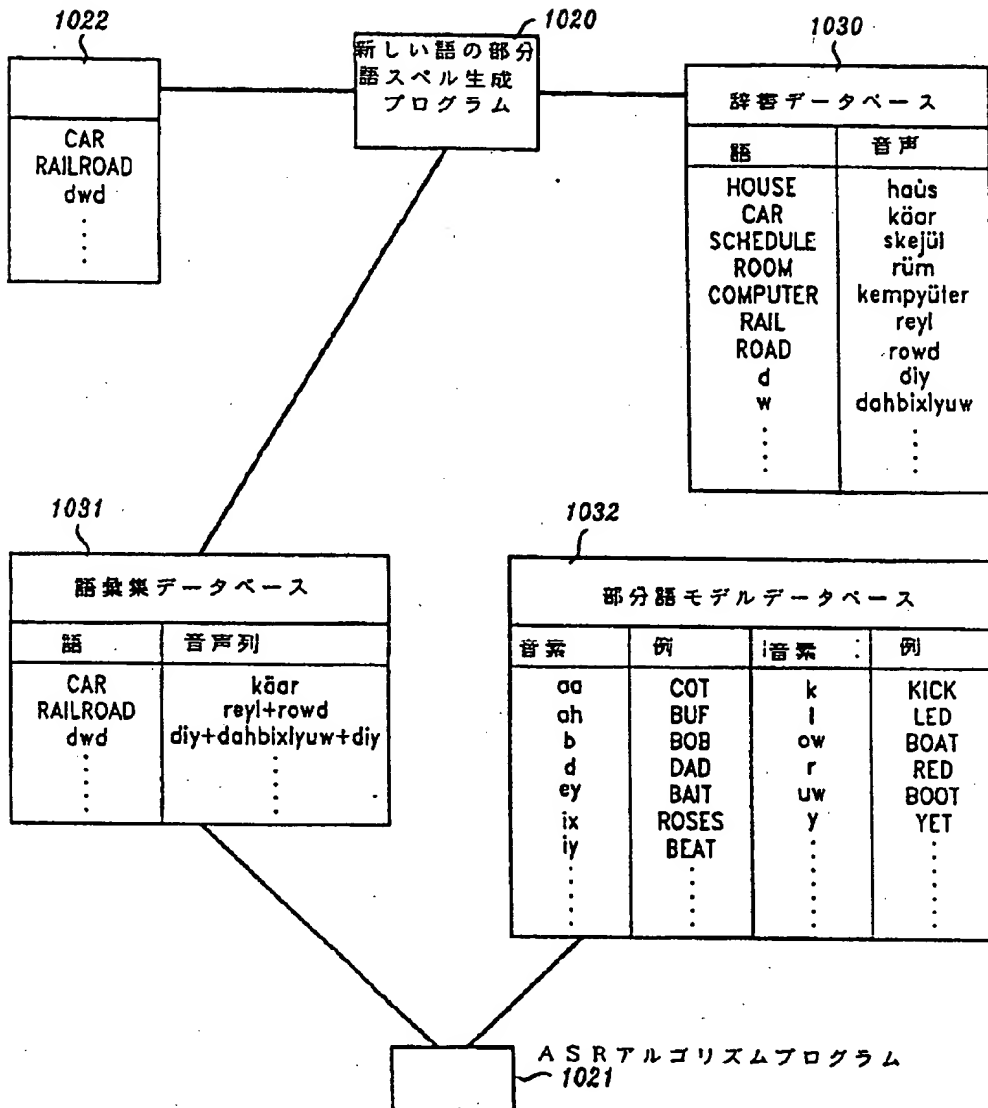


【図2】

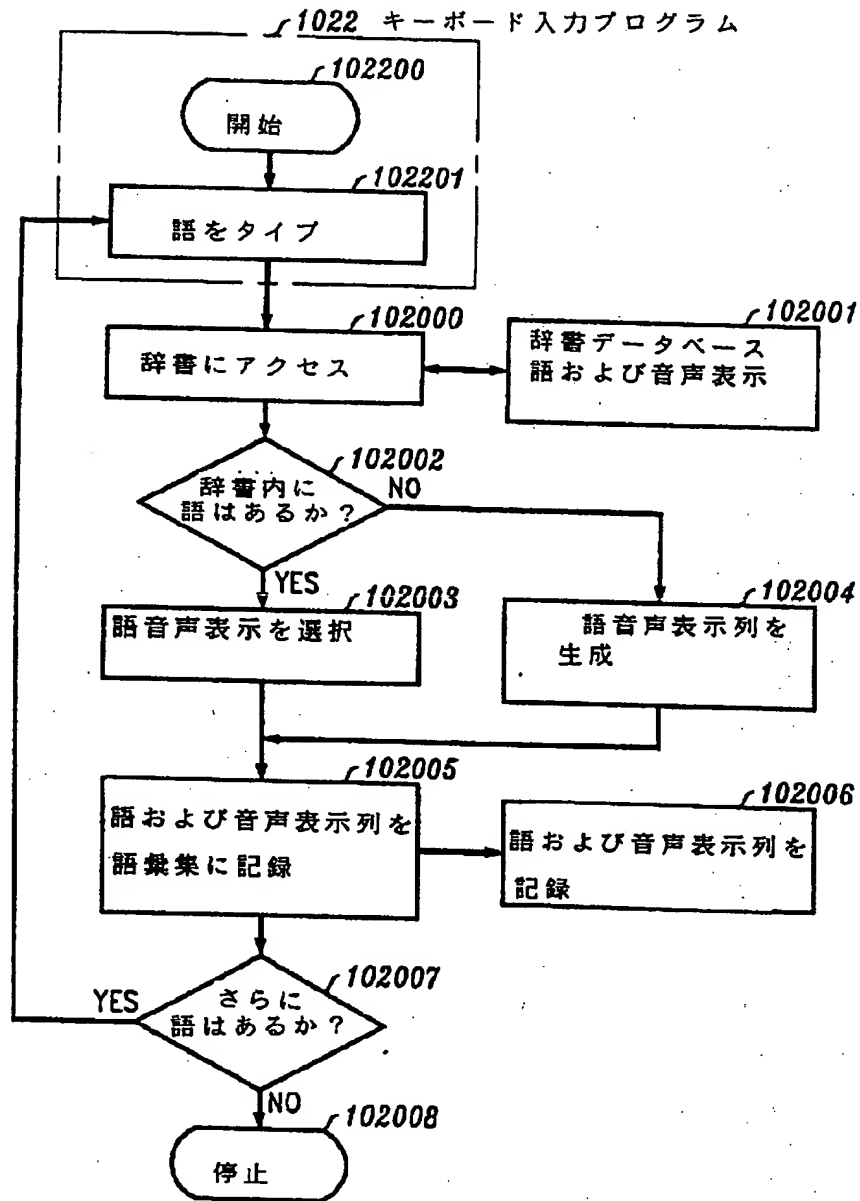


【図3】

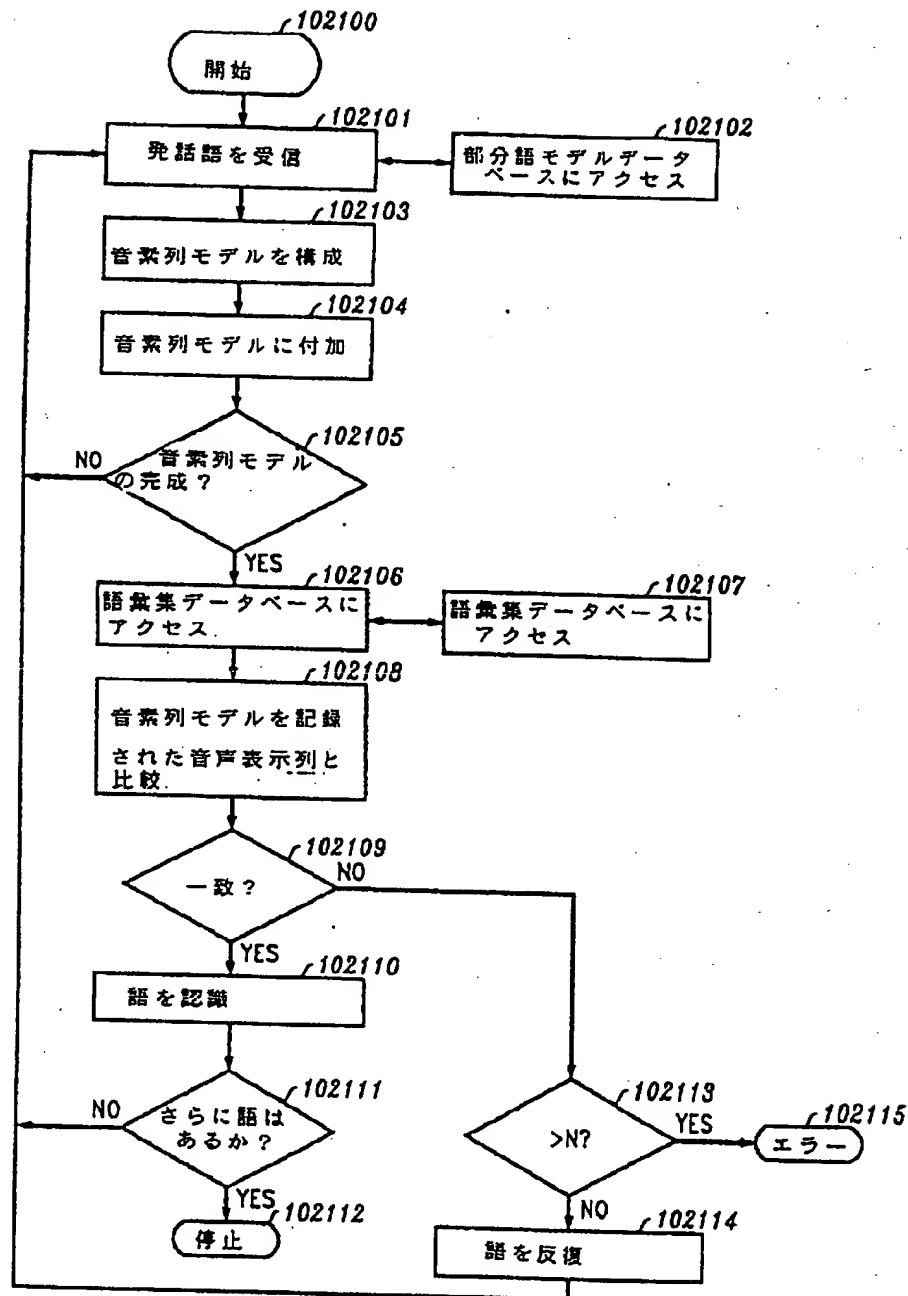
キーボード入力プログラム



【図4】

1020 新しい語の部分語スベル生成プログラム

【図5】



フロントページの続き

(72)発明者 エンリコ エル. ボッチェリ
 アメリカ合衆国 07940 ニュージャージー
 マジソン、ナンバー31 マジソン ア
 ヴェニュー 17

(72)発明者 セダ アイ. ゴクセン
 アメリカ合衆国 43230 オハイオ ガー
 ナ、ウィットレー ドライヴ 474

(72)発明者 ラジェンドラ ビー. ミツキリネニ
アメリカ合衆国 43230 オハイオ ガー
ナ、ステッドウェイ コート 551

(72)発明者 デヴィッド ビー. ロー
アメリカ合衆国 07922 ニュージャージ
ー パークレー ハイツ、オーヴァールッ
ク ドライヴ 8

(72)発明者 ジェイ ゴードン ウィルボン
アメリカ合衆国 07059 ニュージャージ
ー ウォーレン、ラウンド トップ ロー
ド 75